JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2005, **84,** 435–451 NUMBER 3 (NOVEMBER)

RELATING DERIVED RELATIONS AS A MODEL OF ANALOGICAL REASONING: REACTION TIMES AND EVENT-RELATED POTENTIALS

DERMOT BARNES-HOLMES, DONAL REGAN, YVONNE BARNES-HOLMES, SEAN COMMINS, DEREK WALSH, IAN STEWART, PAUL M. SMEETS, ROBERT WHELAN, AND SIMON DYMOND

NATIONAL UNIVERSITY OF IRELAND, MAYNOOTH;
NATIONAL UNIVERSITY OF IRELAND, GALWAY;
LEIDEN UNIVERSITY;
UNIVERSITY COLLEGE, DUBLIN;
AND UNIVERSITY OF WALES, SWANSEA

The current study aimed to test a Relational Frame Theory (RFT) model of analogical reasoning based on the relating of derived same and derived difference relations. Experiment 1 recorded reaction time measures of similar–similar (e.g., ''apple is to orange as dog is to cat'') versus different–different (e.g., ''he is to his brother as chalk is to cheese'') derived relational responding, in both speed-contingent and speed-noncontingent conditions. Experiment 2 examined the event-related potentials (ERPs) associated with these two response patterns. Both experiments showed similar–similar responding to be significantly faster than different–different responding. Experiment 2 revealed significant differences between the waveforms of the two response patterns in the left-hemispheric prefrontal regions; different–different waveforms were significantly more negative than similar–similar waveforms. The behavioral and neurophysiological data support the RFT prediction that, all things being equal, similar–similar responding is relationally ''simpler'' than, and functionally distinct from, different–different analogical responding. The ERP data were fully consistent with findings in the neurocognitive literature on analogy. These findings strengthen the validity of the RFT model of analogical reasoning and supplement the behavior-analytic approach to analogy based on the relating of derived relations.

Key words: relating relations, analogy, reasoning, derived relations, humans

---

Relational Frame Theory (RFT) aims to provide a modern behavior-analytic account of human language and cognition (see Hayes, Barnes-Holmes, & Roche, 2001, for a book-length review). According to the theory, the power, richness, and complexity of human verbal behavior may be usefully interpreted as arbitrarily applicable relational responding. Nonarbitrary examples of relational responding, such as identity and oddity matching, for example, are well established in behavior analysis (e.g., Kastak & Schusterman, 1994). In both types of matching, the relational response is controlled in part by the nonarbitrary or formal relation between the sample and comparison stimuli. Relational frame theory argues that relational responding may be brought under contextual control, in that the relational responses are determined not by the formal properties of the related events but by additional contextual cues. For example, if A is specified as similar to B, then a verbally competent human may state that ''B is similar to A,'' based on the contextual control established by the verbal community for the word *similar*. In this case, the relation of similarity established between A and B is applied arbitrarily (by the verbal community) and is not determined by the physical relation between the stimuli. Arbitrarily applicable relational responses such as this define relational frames.

According to RFT, relational frames possess three defining behavioral properties: mutual entailment, combinatorial entailment, and the transformation of function (Barnes, 1994; Hayes, 1991). Mutual entailment involves establishing, in a given context, a particular relation between Stimulus A and Stimulus B, and then observing the mutually entailed relational response between B and A. The mutually entailed relation depends on the

The authors would like to thank Steve Hayes for his constructive and insightful comments on an earlier draft of the current article.

Dermot Barnes-Holmes, Donal Regan, Y vonne Barnes-Holmes, Sean Commins, and Derek Walsh are at the National University of Ireland, Maynooth; Ian Stewart is at the National University of Ireland, Galway; Paul M. Smeets is at Leiden University; Robert Whelan is at University College, Dublin; and Simon Dymond is at the University of Wales, Swansea.

Correspondence and requests for reprints should be addressed to Dermot Barnes-Holmes, at the Department of Psychology, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland (e-mail: Dermot.Barnes-Holmes@may.ie).

doi: 10.1901/jeab.2005.79-04

type of relation trained between A and B. If Stimulus A bears a *similar* relation to B, for example, then the relation ''B is similar to A'' is entailed (the relation of *similar* is one example of the frame of coordination; Hayes, 1991). Trained and mutually entailed relations are not always identical, however—if responding to A as *taller than* B is trained, for instance, then a *shorter than* relation may be entailed from B to A. The second property of a relational frame, combinatorial entailment, is more complex than mutual entailment. If, for example, Stimulus A bears a relation to B and B bears a relation to C, then a relation between A and C may be derived. Once again, the nature of this derived relation depends on the nature of the trained relations. If A is *taller than* B and B is *taller than* C, for instance, then a *taller than* relation between A and C is derived by combinatorial entailment (i.e., A is taller than C), and a *shorter than* relation is entailed between C and A (i.e., C is shorter than A). Transformation of function, the third property of a relational frame, occurs when a stimulus acquires a particular behavioral function based on its derived relation to another stimulus. If, for example, Stimulus A is related to Stimulus B, and A acquires a specific function, in certain contexts the stimulus functions of B will be transformed in accordance with the A-B relation. If A is the opposite of B, for instance, and A evokes an avoidance response, then B may evoke an approach response (e.g., Whelan & Barnes-Holmes, 2004).

Numerous studies have provided empirical evidence to support the concept of multiple relational frames (e.g., Dymond & Barnes, 1994, 1995, 1996; Steele & Hayes, 1991; Whelan & Barnes-Holmes, 2004). Relational frame theory also makes specific predictions about higher levels of relational complexity in which relational frames are related to relational frames, thereby producing what are referred to as relational networks (e.g., Lipkens, 1992; Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001; Stewart, Barnes-Holmes, Roche, & Smeets, 2001, 2002). In other words, RFT predicts that derived relations may be related to other derived relations (an empirical example of such a complex relational network will be described below). The critical point in the context of the current research is that RFT argues that responding in accordance with relations between relations (hereafter referred to as relating relations) provides the theoretical basis for a functional analysis of the key behavioral properties of analogical reasoning (e.g., Barnes, Hegarty, & Smeets, 1997; Lipkens, 1992; Stewart & Barnes-Holmes, 2001; Stewart et al., 2002).

In the first published study in this area, Barnes et al. (1997) created an empirical model of analogical reasoning based on relating similar relations to similar relations (described here as similar–similar relations) and relating different relations to different relations (termed here different–different relations). To understand the theoretical rationale for this work, consider the following question: ''dog is to cat as apple is to (a) orange, or (b) snail? Assuming that ''dog'' and ''cat'' are deemed similar in the context *animal* and ''apple'' and ''orange'' in the context *fruit,* a language-able person might be expected to pick orange as the correct answer. A technical description of this response in RFT terms would state that two derived relations of similarity were related to each other in accordance with a frame of coordination (see Figure 1, upper panel). In this case, only one *generic* relational frame (coordination) is involved in the three relational responses that constitute the analogical relational network; that is, the two similar relations participate in a frame of coordination because they are coordinate or similar. In contrast, relating difference relations to difference relations seems to parallel a less common analogical structure (e.g., ''he is to his brother as chalk is to cheese''; see Figure 1, lower panel). In this case, two *generic* relational frames (coordination and distinction) are involved in the relational network; that is, the two difference relations participate in a frame of coordination because they are distinct or different.

In the study reported by Barnes et al. (1997), participants were first trained, using a standard matching-to-sample (MTS) procedure, in the following conditional discriminations: A1-B1, A2-B2, A3-B3, A4-B4, A1-C1, A2-C2, A3-C3, and A4-C4. Participants were then tested for the formation of four derived three-element frames of coordination: A1-B1-C1, A2-B2-C2, A3-B3-C3, and A4-B4-C4. Subsequently, during the (critical) relating-relations test phase, participants related compound comparisons to compound samples.

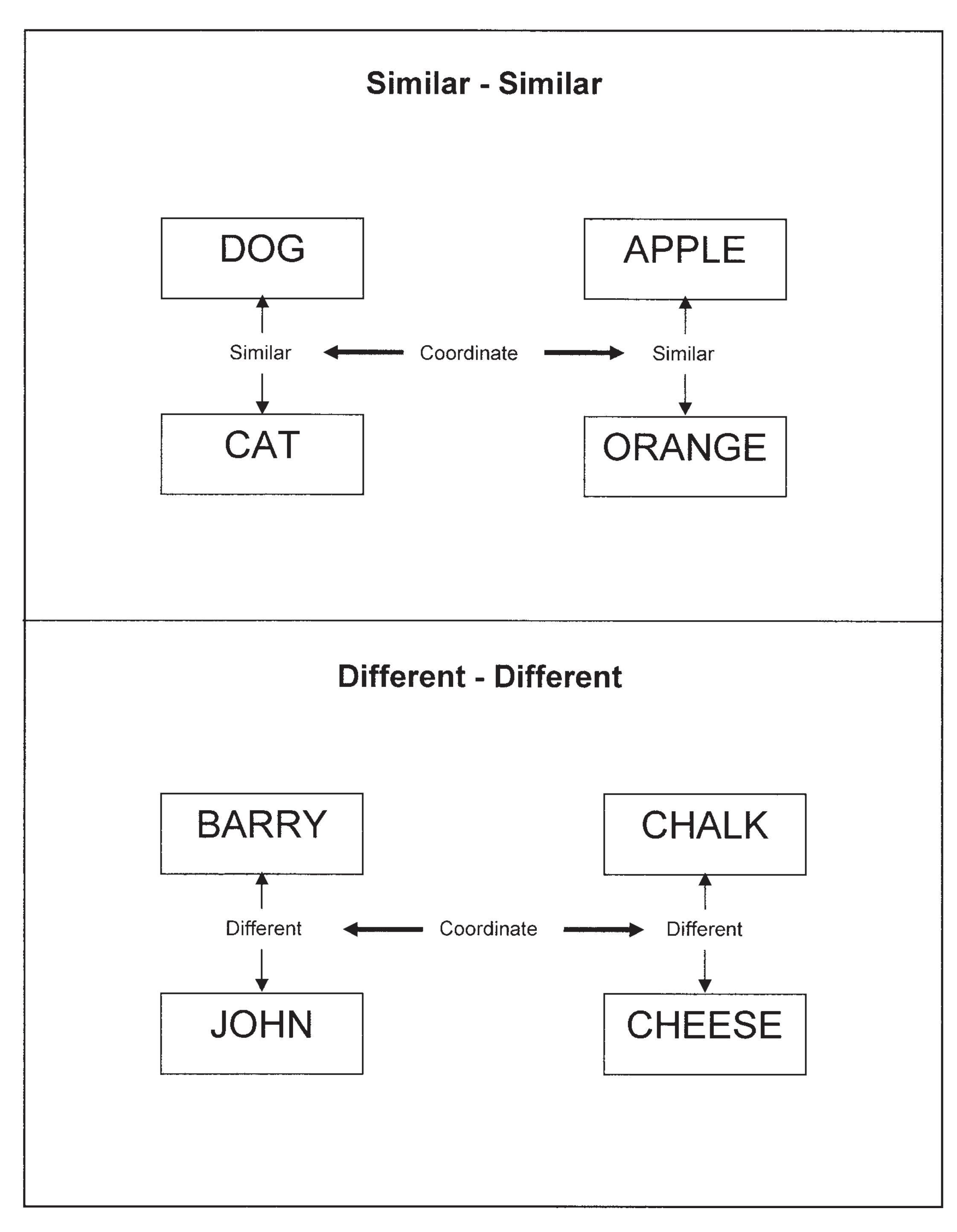


Fig. 1. Schematic representation of the relating of relations. Upper panel: Conceptualizing the analogy ''cat is to dog as apple is to orange'' in terms of the frame of coordination between two derived frames of *similarity/coordination*. Lower panel: Conceptualizing the analogy ''he is to his brother as chalk is to cheese'' in terms of the frame of coordination between two derived *difference* relations.

These compounds comprised two stimulus elements that were either from the same derived relation (e.g., B1C1) and thus were relationally similar, or comprised two elements that were from two different derived relations (e.g., B1C2) and thus were relationally different. All participants, including a 9-year-old boy, demonstrated similar–similar (e.g., B1C1-B3C3) and different–different (e.g., B1C2-B3C4) relational responding. These two patterns of responding provide, according to RFT, models of two different types of analogy.

In conceptualizing analogy in terms of relating derived relations, RFT has generated a growing body of empirical research into various aspects of analogical reasoning in both adults and children (e.g., Carpentier, Smeets, & Barnes-Holmes, 2002, 2003; Carpentier, Smeets, Barnes-Holmes, & Stewart, 2004; Stewart et al., 2002; see Stewart & Barnes-Holmes, 2004, for a review). In all of these studies, however, performance measures always have been restricted to response accuracy (e.g. percentage correct). Reaction time (RT) and neurophysiological measures of relating relations, however, never have been recorded. The current research constitutes a first step towards filling this empirical gap in the literature.

In conducting a chronometric study of the relating of relations, it is reasonable to ask what predictions RFT might make. All things being equal, RT differences between similar–similar and different–different relational responding might be expected because the number of distinct relational frames differs across the two relational networks (see Steele & Hayes, 1991). The different–different analogical network involves placing two different relations in a frame of coordination with each other. That is, two separate relational frames or response classes (distinction and coordination) are being employed to produce the correct relational response. In contrast, the similar–similar analogical network involves placing two frames of coordination in a frame of coordination with each other. This relational network only makes use of one generic relational frame—coordination. One might expect, therefore, that different–different responding will produce longer reaction times than similar–similar responding, because two behavioral classes are required in the former case, but only one is needed in the latter. Testing this prediction was the primary purpose of Experiment 1 (see the General Discussion for an important caveat concerning this prediction).

EXPERIMENT 1

Experiment 1 was a partial replication of Experiment 1 reported by Barnes et al. (1997). The two most important differences were as follows. First, RTs were recorded for all relating-relations test trials. Second, a speed-contingent condition was added (cf. Holth & Arntzen, 2000; Imam, 2001). If correct speed-contingent responding is observed, then the current relating-relations model will be further validated in that natural-language analogical reasoning often occurs rapidly (Klein, 1987). Experiment 2 replicated Experiment 1, but both RTs and electroencephalograms (EEGs) were recorded during the relating-relations test.

Method

Participants

Twenty-four individuals, 10 males and 14 females, agreed to participate. Ages ranged from 18 to 49 years (mean = 24.9). All participants were 1st-year undergraduates at the National University of Ireland, Maynooth. Twelve participants were assigned to the speed-contingent condition and 12 were assigned to the speed noncontingent condition. The majority (91.7%) of participants failed to complete the entire experiment in one session, and were called back on subsequent days. On these days, participants were reexposed to experimental phases that they already had completed (to preserve previously established performances). All participants were asked not to discuss their participation with anyone.

Apparatus and Materials

Standard Dell® PCs (Pentium 4®) with standard 14 in. screens displayed three-letter words on white squares (6 cm by 5 cm), which were set against a light-grey background. Stimulus presentation and the recording of responses, including RTs, were controlled by a custom-made Microsoft® Visual Basic 6.0© training and testing program, written by the first author. Self-adhesive stickers were placed on four response keys on the computer keyboard. The on-screen sample and compar-

ison stimuli were 12 nonsense syllables (FUB, JOM, MAU, DAX, ROG, CUG, PAF, KIB, BEH, YIM, ZID, and VEK) taken from Barnes et al. (1997). For ease of communication in this article, each syllable will be referred to using an alphanumeric label (A1, B1, C1, A2, B2, C2, A3, B3, C3, A4, B4, or C4); participants never saw these labels. For each participant, the nonsense syllables were assigned randomly to their alphanumeric labels.

Delayed Matching-to-Sample

A delayed, arbitrary matching-to-sample (MTS) training and testing procedure was employed. On any given trial, the sample stimulus first appeared center-screen. After the space bar was pressed, the sample disappeared and the screen remained blank for 400 ms. The four comparison stimuli then appeared, one in each corner of the screen. Participants selected one of the comparisons by pressing the corresponding keyboard key: ''R'' for the top left comparison stimulus, ''U'' for top right, ''C'' for bottom left, and ''N'' for the bottom right comparison. The screen then cleared and during training, correct responses were followed immediately by the word ''correct'' in the center of the screen (accompanied by an audible beep). Choosing any of the other three comparisons immediately produced the word ''wrong'' center-screen (with no accompanying sound). Locations of correct/incorrect comparison stimuli were counterbalanced across trials. Feedback (i.e., correct or wrong) remained on screen for 2,000 ms and after a further 2,000 ms intertrial interval (ITI), during which the screen remained blank, a new trial began. During all testing phases, feedback was omitted and the ITI followed directly after a response was made.

During the relating-relations testing phase, sample and comparison stimuli were compounds comprising two stimulus elements presented side-by-side separated by a 1 cm space (e.g., ''BEH DAX''). On each trial, two comparisons were presented after pressing the space bar. These comparisons were placed in the bottom two corners of the screen and were selected by pressing either the C or N keys (all other keys were disabled). Again, comparison location was counterbalanced across trials.

In the speed-contingent condition, participants were required to select a comparison stimulus within 3 s of comparison onset in *all* training and testing trials. If no response was made within this time, the words ''too late!'' appeared center-screen and an incorrect response was recorded. In the speed-noncontingent condition, the computer simply waited until the participant made a response on each trial.

Procedure

Participants were studied individually in a small testing room. At the beginning of each session, participants were given a written instruction sheet that described the matching-to-sample task, including which keys to press to select the comparison stimuli (all verbatim instructions used in both Experiments 1 and 2 are available from the first author upon request). The experimenter, after answering any further queries, then left the room.

Matching-to-Sample Training

Eight MTS tasks were used to train participants in two sets of related conditional discriminations (see Figure 2, top panel). That is, A-B relations: A1-B1, A2-B2, A3-B3, and A4-B4, and A-C relations: A1-C1, A2-C2, A3-C3, and A4-C4. Order of presentation (of the two sets) was counterbalanced across participants. The four tasks of each set were presented in a quasi-random order in four-trial blocks (i.e., each of the four tasks was presented once within each block), until participants produced a minimum of eight consecutively correct responses. After training to criterion on both the A-B and A-C relations, all eight tasks were then presented in a quasi-random order in eight-trial blocks (each of the eight tasks was presented once within each block), until participants produced a minimum of 16 consecutively correct responses. Participants then proceeded directly to the derived relations test.

Derived Relations Test

At the beginning of this phase, a message appeared on the computer screen informing participants that the computer would no longer indicate if their responses were correct or wrong. Twenty-four different MTS tasks were used during this phase of the experiment; eight of these tested for the directly trained relations described above, eight tested for the derived mutually entailed relations

Matching-to-Sample Training

B4 B3 / **A1** / **B1** B2	B4 B3 / **A2** / B1 **B2**	B4 **B3** / **A3** / B1 B2	**B4** B3 / **A4** / B1 B2
C4 C3 / **A1** / **C1** C2	C4 C3 / **A2** / C1 **C2**	C4 **C3** / **A3** / C1 C2	**C4** C3 / **A4** / C1 C2

Derived Mutually Entailed Relations

A4 A3 / **B1** / **A1** A2	A4 A3 / **B2** / A1 **A2**	A4 **A3** / **B3** / A1 A2	**A4** A3 / **B4** / A1 A2
A4 A3 / **C1** / **A1** A2	A4 A3 / **C2** / A1 **A2**	A4 **A3** / **C3** / A1 A2	**A4** A3 / **C4** / A1 A2

Derived Combinatorially Entailed Relations

B4 B3 / **C1** / **B1** B2	B4 B3 / **C2** / B1 B2	B4 B3 / **C3** / B1 B2	B4 B3 / **C4** / B1 B2
C4 C3 / **B1** / **C1** C2	C4 C3 / **B2** / C1 **C2**	C4 **C3** / **B3** / C1 C2	**C4** C3 / **B4** / C1 C2

Relating Derived Relations

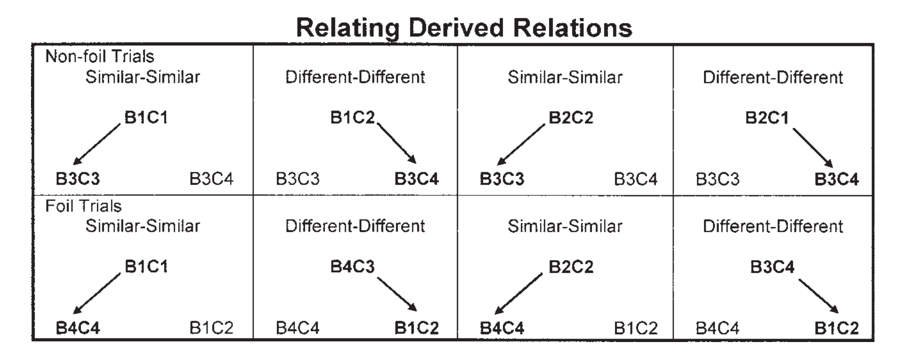


Fig. 2. A schematic representation of the matching-to-sample tasks, including correct responses, which were used in the three phases of Experiment 1: matching-to-sample training, derived relations test, and relating-relations test.

(i.e., four B-A and four C-A relations; Figure 2, second panel), and the final eight tasks tested for the derived combinatorially entailed relations (four B-C and four C-B relations; Figure 2, third panel). The 24 tasks were presented, without feedback, in a random order in a single 24-trial block (i.e., each task was presented once within a block). Participants were required to produce at least 23 of 24 correct responses to pass the derived relations test. Failure to reach this criterion resulted in the message: ''You may take a break now'' appearing center-screen. When participants clicked on a ''continue'' button with the mouse they were returned to the MTS training phase. Following successful retraining, participants were reexposed to the derived relations test, and if necessary, were retrained and retested until they met the pass criterion (≥ 23 of 24 correct responses). Upon reaching the pass criterion, participants proceeded directly to the relating-relations test.

Relating-Relations Test

At the beginning of this phase, a message appeared on the computer screen informing participants that only two choices would be presented on each trial and that each stimulus box would contain two nonsense words (rather than one). Participants were instructed to rest their hands comfortably at the bottom of the keyboard and to place the left and right index fingers on the marked C and N keys, respectively, and to use the right thumb to press the space bar (to remove the sample stimulus). Participants also were instructed to press the left key to choose the left comparison and the right key to choose the right comparison; the other two marked keys, R and U, were inactive throughout the relating-relations test.

Participants were presented with one pair of nonsense syllables as a sample and two pairs of nonsense syllables as comparisons. Each pair of syllables was either from the same derived relation (e.g., B1C1) or from two different derived relations (e.g., B2C1). The relating-relations test consisted of a single block of 32 trials during which eight different MTS tasks (see Figure 2, bottom panel) were presented in a quasi-random order such that each task was presented four times. Consistent with Barnes et al. (1997), four of the eight tasks were labelled ''foils,'' defined as trials in which incorrect comparisons contained an element that also was present in the sample. The sample stimuli were B1C1, B2C2, B1C2, and B2C1 (for the four nonfoil tasks) and B1C1, B2C2, B4C3, and B3C4 (for the four foil tasks). The comparison stimuli for the four nonfoil tasks were B3C3 and B3C4, and the comparison stimuli for the four foil tasks were B4C4 and B1C2.

Relating derived relations was defined as follows: (a) given a sample that contained two elements from the same derived relation (e.g., B1C1), the correct comparison contained two elements from another same relation (e.g., B3C3); and (b) given a sample that contained two elements from different derived relations (e.g., B1C2), the correct comparison contained two elements from another two different derived relations (e.g., B3C4). Two of the eight relating-relations tasks were labelled ''similar–similar,'' two ''different–different,'' another two ''similar–similar with foil,'' and the final two ''different–different with foil'' (see Figure 2, bottom panel).

To pass the relating-relations test, a minimum of 29 of 32 predicted responses were required. At the end of a 32-trial block, regardless of performance, a message appeared center-screen asking the participant to report to the experimenter. If the prearranged session-time had not fully elapsed at this stage, and the participant had failed to reach criterion, she or he was cycled through the entire procedure again, beginning with matching-to-sample training. If the participant had failed, and the session-time had elapsed, a subsequent session-time was arranged. If the participant had passed, Experiment 1 was complete for this individual.

Results and Discussion

Training and Testing Exposures

The mean number of training and testing cycles required to pass the relating-relations test was 5.8 in the speed-contingent condition and 3.5 in the speed-noncontingent condition; this difference approached, but did not reach significance, $F(1, 22) = 3.81$, $p = .064$.

Relating-Relations Test: Reaction Time Measures

Individual RTs were measured from the point of onset of the two comparison stimuli to the participant's first key press (either C or N). Only RTs for experimenter-defined correct responses from *passed* relating-relations tests are

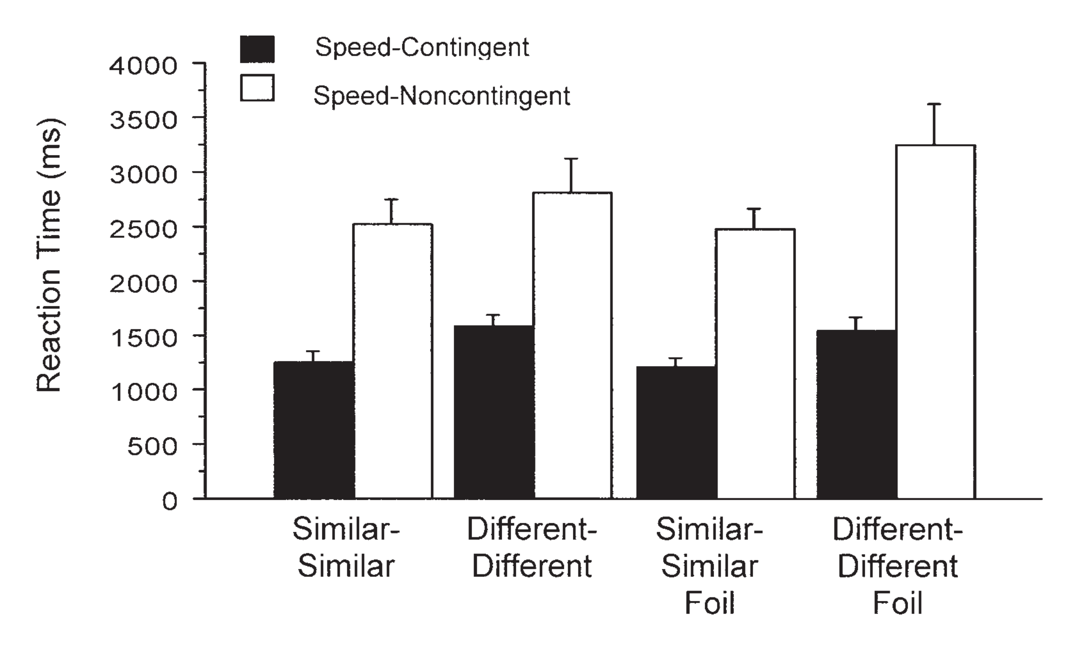


Fig. 3. Mean reaction times (in ms), including standard error bars, for the four relating-relations trial types as a function of experimental condition in Experiment 1.

presented in the current report. Responses of 10 s or over in the speed-noncontingent condition were removed as outliers.

The mean RTs and standard errors for each of the four relating-relations trial types (similar–similar, different–different, similar–similar with foil, and different–different with foil) were calculated across the 24 participants in the two speed conditions (see Figure 3). Mean RTs in the speed-noncontingent condition approached or exceeded twice the duration of the RTs in the speed-contingent condition, across all four trial types. Furthermore, mean RTs for similar–similar trial types were shorter than for different–different trial types in both speed conditions and for both foil and nonfoil trials.

A mixed 2 × 4 analysis of variance (ANOVA) was conducted to examine the effect of the speed contingency (between-participant variable), the effect of the four trial types (within-participant variable), and any possible interaction effect between these two variables. A statistically significant main effect for speed condition, $F(1, 22) = 29.21$, $p <.001$, $\eta_p^2 = .57$, and trial type, $F(3, 66) = 7.1$, $p < .001$, $\eta_p^2 = .24$, was found but no significant interaction effect was detected, $F(3, 66) = 1.3$, $p = .278$. A series of Fisher's PLSD post hoc tests were then conducted, comparing RTs across the four relating-relations trial types (see Table 1). Similar–similar responses, in both speed conditions, were significantly faster than different–different responses. The RTs for the foil trials did not differ significantly from those of their nonfoil counterparts.

Relating-Relations: Accuracy Measures

In addition to RT measures of relating-relations trial types, *accuracy* measures also were compared. Because the accuracy of these four trial types would be almost identical for

Table 1

Results of Fisher's PLSD post hoc tests comparing reaction times for the four relating-relations trial types in both speed conditions of Experiment 1, with percentage of individual participants conforming to the direction of the statistically significant effects.

Post hoc comparisons	Significance (*p*)	%
Similar–similar vs. different–different	.027	83.3
Similar–similar vs. similar–similar (foil)	.744	
Similar–similar vs. different–different (foil)	.001	83.3
Different–different vs. similar–similar (foil)	.012	75.0
Different–different vs. different–different (foil)	.176	
Similar–similar (foil) vs. different–different (foil)	.0002	83.3

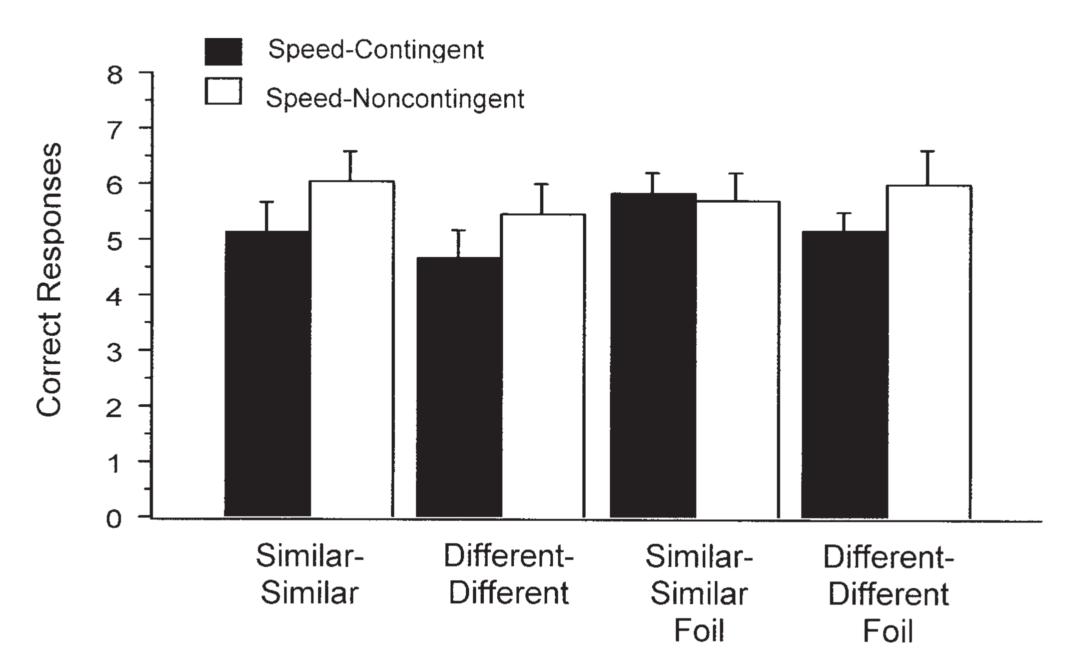


Fig. 4. Acquisition data averaged across 21 participants showing the total or mean number of correct responses (out of eight), including standard error bars, across the four relating-relations trial types, as a function of experimental condition in Experiment 1.

the final (i.e., passed) exposure to the relating-relations test, only participants' *acquisition* data (from relating-relations tests prior to the final successful test) were analyzed. Twenty-one of the 24 participants failed at least one relating-relations test before completing the experiment, and thus their acquisition data could be analyzed (the three remaining participants who passed the first relating relations test were from the speed-noncontingent condition). For the analysis, four values were calculated for each participant: the total number of correct responses (out of eight) made during similar–similar and different–different test trials with and without foils. If a participant failed more than one relating-relations test, the mean (rather than the total) number of correct responses across the failed tests was calculated for each of the four conditions. The accuracy data averaged across the 21 participants is shown in Figure 4. Apart from one trial type (similar–similar with foil), imposing the speed-contingency appeared to affect accuracy negatively, but only slightly. No other consistent and obvious trial type effects emerged for the accuracy measure.

A mixed 2 × 4 ANOVA was conducted to examine the effect of the speed contingency, the effect of the four trial types, and a possible interaction between these two variables. The speed contingency did not reach significance, $F(1, 19) = 1.19$, $p = .290$, nor did trial type, $F(3, 57) = 1.47$, $p = .234$, nor was any interaction between these variables detected, $F(3, 57) = .93$, $p = .434$. In effect, the accuracy measure was not affected, at a statistically significant level, by either the speed-contingency or the relating-relations trial type.

Summary

Experiment 1 produced three clear effects. First, the results showed that average RTs for different–different trials were significantly longer than for similar–similar trials—this difference was not affected by imposing a speed-contingency, nor was it affected by the foils. Second, adding the speed-contingency significantly reduced RTs across all four trial types. Third, accuracy measures derived from the relating-relations acquisition data were not significantly affected by speed condition or trial type.

The results of Experiment 1 provide chronometric evidence for the RFT model of analogical reasoning in that longer RTs were recorded when participants were required to relate derived relations that involved two generic classes of relational responding (Same

and Different), rather than only one (Same). There are, however, other measures that might be used to investigate the relating of derived relations as a model of analogical reasoning. One of these involves the use of EEG signals in the form of event-related potentials. The primary purpose of Experiment 2 was to commence such work.

EXPERIMENT 2

Due to recent technological advances in the measurement of brain events, it now is possible to study verbal processes and hemispheric specialization with greater precision than can be accomplished with behavioral measures alone (Andreassi, 2000). Event-related potentials (ERPs) constitute one such technological innovation (see Barnes-Holmes et al., 2005). The ERP measure is constituted from averaged EEG segments, time-locked to some repeatedly-presented stimulus event, and the signals are typically recorded from various positions around the scalp.

Although a literature search failed to find any ERP studies investigating analogical reasoning, alternative neuroscience methodologies have been employed in this area. For example, a recent positron emission tomography (PET) study found left-hemispheric neural activation while participants engaged in a figural analogy task (Wharton et al., 2000). This activity was specifically located in the left dorsolateral prefrontal cortex (PFC). Wharton et al. found little or no accompanying right-hemispheric activity and so concluded that even figural analogies are processed linguistically in the left PFC. A later study examined the effects of repetitive transcranial magnetic stimulation (rTMS) of the left and right dorsolateral PFC on figural analogical reasoning in 16 normal adults (Boroojerdi, Phipps, Kopylev, Wharton, Cohen, & Grafman, 2001). Stimulation over the left PFC led to significantly reduced reaction times in the analogy task; right PFC stimulation did not. The authors concluded that left PFC is relevant for analogical reasoning.

Another recent study in the area (Luo et al., 2003), used functional magnetic resonance imaging (fMRI) in the analysis of verbal analogy, and found *bilateral* hemispheric activation. Specifically, the left and right PFCs were activated, as well as bilateral temporal structures such as the fusiform gyri and basal ganglia. Because some analogies in Luo et al's. study (although all verbal) may have been more visually than linguistically based, the authors suggested that the observed bilateral activation was not entirely surprising.

Perhaps one of the reasons for the dearth of ERP research on verbal analogical reasoning is the lack of a precise methodology for studying such behavior in this way. The interpretation of ERP data generated by analogies embedded in sentences, for example, may be difficult because of confounds such as syntax, context, and a heavy load on semantic working memory (e.g., Weisbrod et al., 1999). Insofar as analogical reasoning, as well as sentence processing, is heavily dependent on working memory, a difficulty would arise in separating out these two activities in an ERP analysis. Furthermore, words used in these sentences would of course carry preestablished semantic associations that would be difficult to control for.

The model of analogy reported in Experiment 1, however, lends itself quite well to an ERP analysis for the following reasons. First, the relating-relations test for analogical responding involves *discrete stimulus presentations*, the effects of which ERPs are specifically designed to measure (Holcomb, 1988; Kutas, 1993). Furthermore, because the current model involves isolated relational responses, confounding variables such as syntax, context, and a heavy load on working memory are minimized. Finally, the stimuli used in the relating-relations model are arbitrary nonsense syllables, and thus the probability of confounding preexperimental semantic histories is much reduced (see Barnes-Holmes et al., 2005).

The findings of Experiment 1 suggest that similar–similar and different–different relational responding are functionally distinct. Given the RT difference between the relating of similar versus difference relations, it would seem reasonable to expect that such differences also would be reflected at the neurophysiological level. Experiment 2 was a partial replication of Experiment 1 and focused on measuring the electrophysiological correlates (ERPs) of similar–similar versus different–different analogical responding. Only the speed-contingent condition was employed, thus providing a fixed and discreet time-window for the ERP analyses. Because previous

PET, rTMS, and fMRI studies (mentioned above) have shown analogical reasoning (both verbal and figural) to be associated with the frontal regions of the cortex, Experiment 2 measured ERPs from six of the standard frontal sites located on the left and right hemispheres.

Method

Participants

Fourteen right-handed individuals, 5 males and 9 females, attending the National University of Ireland, Maynooth, completed the experiment. Ages ranged from 18 to 49 years with a mean age of 23 years. None of the participants had a history of psychiatric or neurological disorder.

Apparatus and Materials

The experiment was conducted in a small, dimly lit laboratory in the Department of Psychology at the National University of Ireland, Maynooth. The stimuli and materials used were identical to those of Experiment 1. In order to record the EEG measures during the relating-relations task, a Brain Amp MR (Class IIa, Type BF) with approved control software (Brain Vision Recorder 1.0), and electrode cap (BrainCap/BrainCap MR) was used. The Brain Amp was controlled by a Dell® personal computer with a Pentium 4® processor. The ERP data were analyzed using approved analysis software (Brain Vision Analyzer 1.0). Hardware and software were manufactured and supplied by Brain Products GmbH, Munich, Germany.

Pretraining and Pretesting

The procedure was identical to that of the speed-contingent condition in Experiment 1. To ensure accurate performance on the day of recording, and to reduce participant fatigue during ERP recording, all participants underwent MTS training and derived relations testing across one or more days immediately prior to the ERP recording phase.

Training and Testing

On the day of ERP measurement, participants first were attached to the Brain Amp and then were reexposed to the MTS training and derived relations testing before proceeding for the first time to the relating-relations test. If necessary, participants were cycled and recycled through the entire three-stage procedure until they passed the relating-relations test. The entire session, including electrode placement, lasted on average 1 hr and 30 min. Only the ERP data from the final successful exposure to the relating-relations test were analyzed.

Recordings

Evoked potentials were recorded and analyzed from six sintered AG/AG-CI scalp electrodes positioned according to the international 10-20 system. The primary purpose of the ERP recordings was to determine if differential EEG patterns would be observed between the two relating-relations trial types in the frontal cortex. The six frontal sites chosen for recording were Fp1, Fp2, F3, F4, F7, and F8. The central vertex electrode was used as reference and the FPz as ground. Two additional electrodes recorded vertical and horizontal eye movements. Amplifier resolution was 0.1 μV (range, ± 3.2768 mV) and the bandwidth set between 0.5 and 62.5 Hz with sampling rate of 250 Hz. The notch filter was set at 50 Hz. All electrode impedances were at or below 5 KΩ. The EEG was collected continuously and edited off-line. Average ERPs were calculated across the similar–similar and different–different test trials.

Results and Discussion

The mean RTs and standard errors for correct responses on the final (successful) exposure to the relating-relations test, for each of the four trial types, were calculated across 13 participants (see Figure 5); the RT data for one of the participants was lost due to experimenter error. Consistent with the results of Experiment 1, mean RTs for similar–similar trial types were lower than for different–different trial types, and the foils appeared to have no effect on RT. A one-way repeated measures ANOVA was conducted to compare RTs across the four trial types, and this proved to be significant, $F(3, 36) = 5.54$, $p = .003$, $\eta_p^2 = .32$. A series of Fisher's PLSD post hoc tests were then conducted, comparing RTs among the four relating-relations trial types (see Table 2). These tests indicated that similar–similar responses were significantly faster than different–different responses, but foil trials did

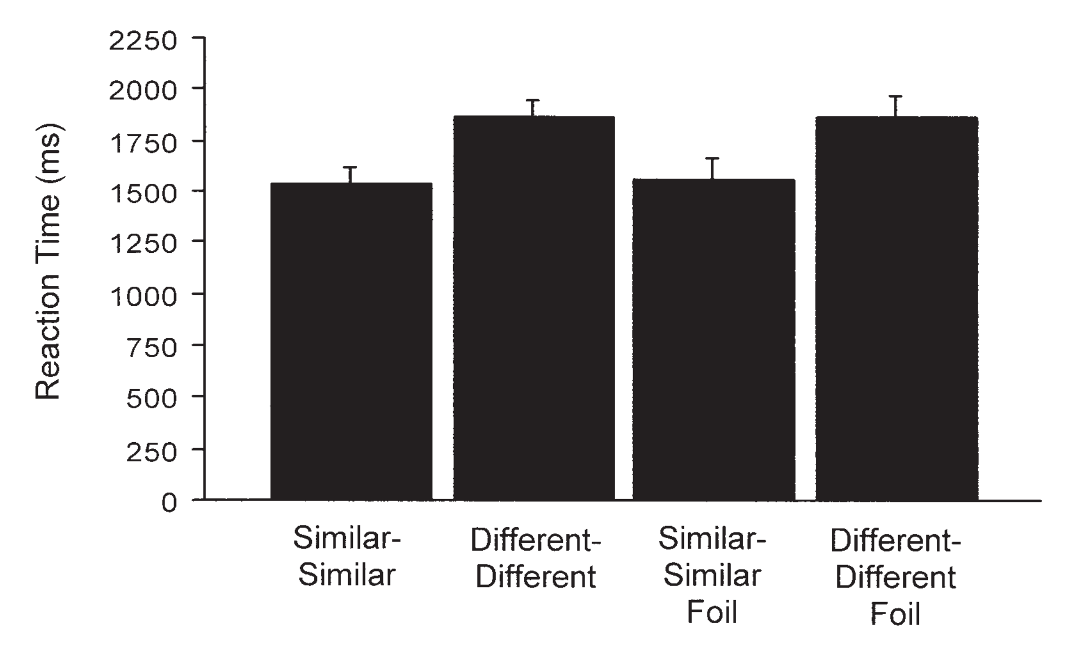


Fig. 5. Mean reaction times (in ms), including standard error bars, for the four relating-relations trial types in Experiment 2.

not differ significantly from their nonfoil counterparts.

The continuous EEG signals for each of the 14 participants were filtered (0.53 Hz, time constant = 0.3 s, 24 dB/octave roll-off) and then segmented for similar–similar and different–different trial types. Standard and foil trials were collapsed to reduce noise for the ERP analysis. The segments were divided into 1,700 ms epochs commencing 100 ms before onset of the comparison stimuli. Vertical and horizontal ocular artifacts were then corrected, and any segments on which EEG or electro-ocular activity exceeded ± 75 μV were rejected. The remaining segments then were baseline corrected (using the 100 ms prestimulus interval) and finally averaged for similar–similar and different–different trial types.

Table 2

Results of Fisher's PLSD post hoc tests comparing reaction times for the four relating-relations trial types in Experiment 2, with percentage of individual participants conforming to the direction of the statistically significant effects.

Post hoc comparisons	Significance (*p*)	%
Similar–similar vs. different–different	.005	84.6
Similar–similar vs. similar–similar (foil)	.819	
Similar–similar vs. different–different (foil)	.005	84.6
Different–different vs. similar–similar (foil)	.009	76.9
Different–different vs. different–different (foil)	.976	
Similar–similar (foil) vs. different–different (foil)	.009	76.9

The grand average waveforms for each of the six frontal electrode sites (Fp1, Fp2, F3, F4, F7, and F8) for similar–similar (light lines) versus different–different trials (dark lines) are shown in Figure 6. Visual inspection of these waveforms indicated little evidence of differential activity for the two trial types until approximately 1,000 ms following comparison onset. Specifically, the different–different trials, relative to the similar–similar trials, produced greater negativity on the left sites (left panels) and more positivity on the right sites (right panels). This differential between trial types was more evident for the left than for the right sites, and the largest difference was observed at the dorsolateral prefrontal regions (F7 and F8).

For the purposes of statistical analysis, the ERP waveforms were divided into three temporal regions: 200 to 600 ms, 600 to 1,000 ms, and 1,000 to 1,400 ms. The area dimensions (μV × ms) for each waveform for each

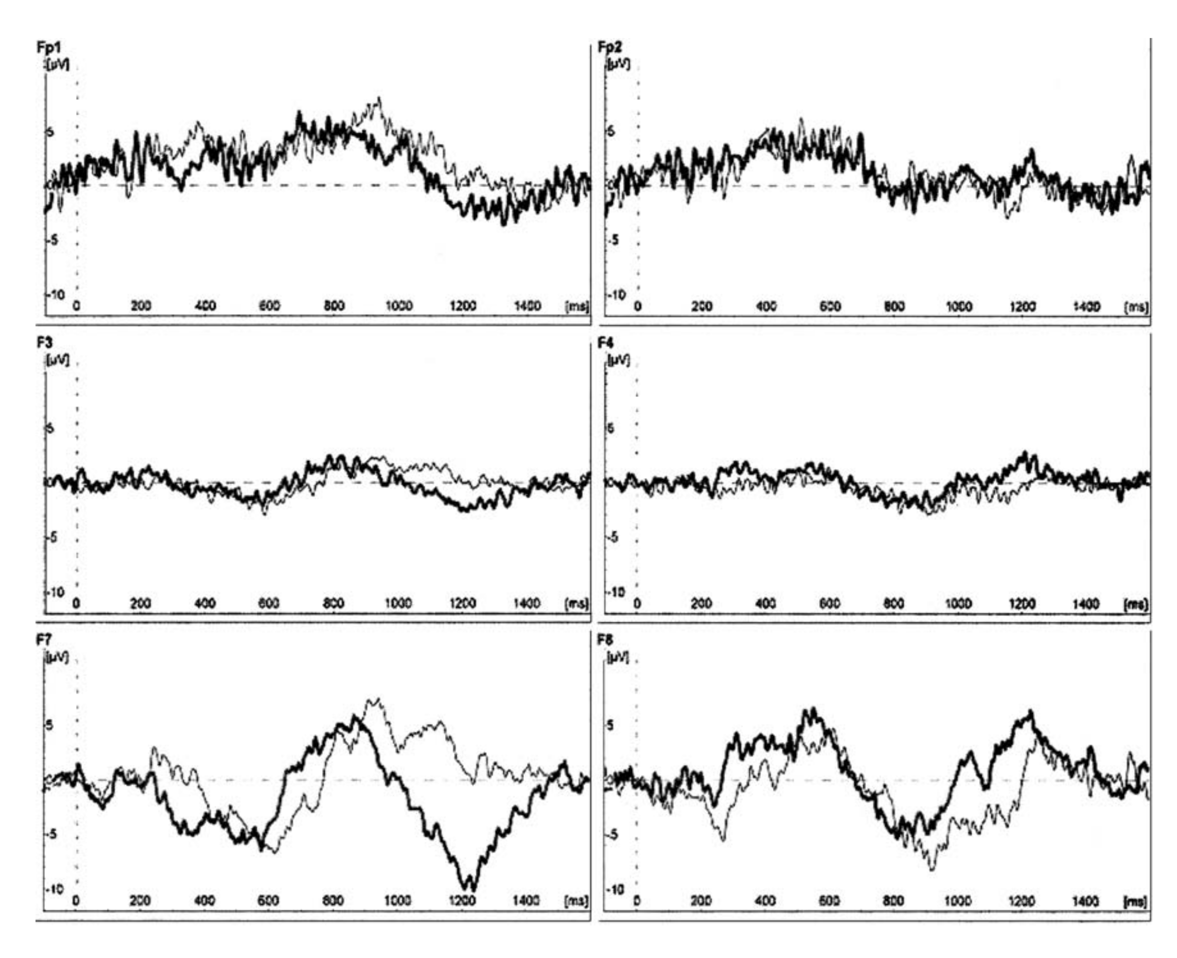


Fig. 6. Grand average waveforms for similar–similar (thin lines) and different–different (thick lines) derived relational responding at electrode sites Fp1, Fp2 (top panel), F3, F4 (middle panel), and F7, F8 (bottom panel). Comparison compounds were presented at 0 ms.

participant recorded during these regions were calculated, yielding either positive or negative values with respect to the 0 µV level. A four-way ANOVA was then conducted with laterality (left and right), location (Fp1-Fp2, F3-F4, F7-F8), interval (200 to 600 ms, 600 to 1,000 ms, 1,000 to 1,400 ms), and trial type (similar–similar and different–different) as repeated measures factors. The ANOVA revealed significant main effects for location, $F(2, 26) = 7.75$, $p = .002$, $\eta_p^2 = 0.37$, and interval, $F(2, 26) = 10.82$, $p < .001$, $\eta_p^2 = 0.45$, but not for laterality, $F(1, 13) = .002$, $p = .962$, or trial type, $F(1, 13) = .786$, $p = .392$. Interaction effects were also found: interval by location, $F(4, 52) = 22.34$, $p = <.001$, $\eta_p^2 = 0.63$, and interval by trial type, $F(2, 26) = 7.05$, $p = .004$, $\eta_p^2 = 0.35$, were both statistically significant.

Given the significant intervallic interactions, three 3-way repeated measures ANOVAs were required to analyse each interval separately. In the first (200 to 600 ms) and second (600 to 1,000 ms) intervals, location was the only significant effect, $F(2, 26) = 14.04$, $p < .001$, $\eta_p^2 = .52$ and $F(2, 26) = 11.64$, $p < .001$, $\eta_p^2 = .47$, respectively. For both intervals, Sheffe post hoc tests then were conducted to compare the three frontal regions of both hemispheres. The tests showed that the Fp1-Fp2 regions differed significantly from both the F3-F4 (first interval, $p < .001$; second interval, $p < .002$) and F7-F8 (first interval, $p < .001$; second interval, $p = .001$) regions but that the F3-F4 and F7-F8 regions did not differ significantly (first interval, $p = .862$; second interval, $p = .992$). To summarize, the waveforms between 200 and 1,000 ms for the Fp1 and Fp2 sites were

significantly more positive than for the F3-F4 and F7-F8 sites.

The third interval (1,000 to 1,400 ms) was not significant for location, $F(2, 26) = .36$, $p = .700$, but did reveal a significant main effect for trial type, $F(1, 13) = 8.45$, $p = .012$, $\eta_p^2 = .39$. There also were two significant interactions: laterality by trial type, $F(1, 13) = 8.84$, $p = .011$, $\eta_p^2 = .40$, and laterality by location by trial type, $F(2, 26) = 8.80$, $p = .001$, $\eta_p^2 = .40$. Two 2-way repeated measures ANOVAs thus were used to explore location and trial type effects in the right and left hemispheric final intervals, respectively. The right-hemispheric ANOVA approached, but did not reach significance, $F(1, 13) = 4.29$, $p = .059$, $\eta_p^2 = .25$. In the left hemisphere, trial type *was* significant, $F(1, 13) = 11.32$, $p = .005$, $\eta_p^2 = .47$, and it also interacted with location, $F(2, 26) = 5.59$, $p < .01$, $\eta_p^2 = .30$. To investigate this interaction further, three separate analyses were conducted for each left-hand site. The three 1-way repeated measures ANOVAs were all significant for trial type: Fp1, $F(1, 13) = 4.93$, $p < .05$, $\eta_p^2 = .27$; F3, $F(1, 13) = 8.72$, $p = .011$, $\eta_p^2 = .40$; and F7, $F(1, 13) = 9.18$, $p < .01$, $\eta_p^2 = .41$. That is, different–different analogical responses, in the left hemisphere and for the final interval, were significantly more negative than their accompanying similar–similar responses at all three left-frontal sites. The percentage of individual participants conforming to the direction of the statistically significant effects for each of the three sites was as follows: Fp1 = 57.1; F3 = 85.7; F7 = 85.7.

The results of Experiment 2 showed that, in the context of a speed-contingency, similar–similar responses were faster than different–different responses, and furthermore that this difference was not affected by foils. The ERP analyses indicated overall differences between waveforms in the three time intervals as well as overall differences between waveforms in the Fp1-Fp2, F3-F4, and F7-F8 areas; moving in a roughly anterior-posterior direction across the prefrontal cortex, slow-wave amplitudes increased and waveforms typically became more negative. Significant differences between similar–similar versus different–different responses were found for the left hemisphere beginning around 1,000 ms after comparison onset. This difference increased as a function of electrode location; moving in an anterior-posterior direction (i.e., from Fp1 to F3 and then F7) different–different waveforms became increasingly more negative than similar–similar waveforms. Interestingly, although only approaching significance, the same pattern was observed for the right hemisphere, except that the differential relation reversed; similar–similar waveforms were slightly more negative than different–different waveforms.

GENERAL DISCUSSION

In both Experiments 1 and 2, significant RT differences were found between the two relational networks (i.e., similar–similar versus different–different). Furthermore, this effect was observed within a compressed 3-s response window. These findings confirm the predictions of RFT (Hayes et al., 2001). That is, all things being equal, a different–different relational network should generate longer RTs than a similar–similar network because the former involves responding in accordance with a similar relation between two difference relations (i.e., *two* distinct relational frames are involved: distinction and coordination); in contrast, similar–similar responses involve responding in accordance with a similar relation between two similar relations (i.e., only *one* relational frame is involved: coordination).

Although the current data support the RFT prediction that multiple stimulus relations will produce longer RTs than single stimulus relations (see Steele & Hayes, 1991), it is important to recognize that RFT is an operant theory and thus its predictions always are constrained by historical and contextual variables. Recognizing this fact is particularly important in the uncontrolled world of natural language where an individual's history of derived relational responding typically is unknown. Consider, for example, the analogy ''midget is to giant as mouse is to elephant // apple is to baseball.'' According to RFT, this analogy involves two comparative relations (relative size) and a similar relation between the two comparatives (see Lipkens, 1992), and thus one might predict that it should produce longer RTs than a similar–similar analogy. Without knowing an individual's history of responding in accordance with such relational networks, however, this prediction cannot be made. It may be, for example, that such comparative analogies are at a relatively high

frequency in natural language and thus occur rapidly for most members of the verbal community. In this case, RT may be rendered a relatively blunt instrument for separating out different analogical networks.

The issue is further complicated because RFT also would predict the current data based on the history that was provided within the experimental context itself. Specifically, the participants were required to pass a test for the frame of coordination before proceeding to the relating-relations test, and thus the operant class of coordination, relative to distinction, may have been at a higher strength when participants entered the relating-relations test. Perhaps, therefore, subsequent exposures to the relating-relations test would have removed the RT difference between similar–similar and different–different relational networks (see Barnes-Holmes et al., 2004).

It is these very concerns about RT, as a measure of derived relational responding, that highlights the importance of exploring alternative dependent variables such as ERP. Indeed, preliminary evidence suggests that ERPs may remain relatively sensitive to different derived relational networks when RT does not (Barnes-Holmes et al., 2004). Furthermore, ERPs and indeed other neurophysiological measures can be used to determine if RFT models of human language and cognition produce effects broadly similar to those seen in neurocognitive research (see Barnes-Holmes et al., 2005). If some degree of overlap is found, then RFT as a behavioral theory of human language and cognition will be strengthened. Certainly, the data obtained in Experiment 2 of the current study appear supportive.

Experiment 2 showed electrophysiological activity associated with both similar–similar and different–different relational responding primarily in the dorsolateral prefrontal areas (F7 and F8). Furthermore, it was in these areas that the greatest differentials between relating similar and difference relations were observed (although only the left hemisphere proved to be statistically significant). Interestingly, differential waveforms in these areas only began to emerge around 1,000 ms *after* presentation of the two comparison stimuli. It seems likely that for the first 1,000 ms, participants were viewing and evaluating *both* (i.e., the correct and incorrect) comparisons independently; that is, without a view to immediately matching them to the sample. Then, at around 1,000 ms, a reliable slow-wave negative deflection for F7 (i.e., an N1000 component) was observed for different–different analogical responses. Given the mean RT for this response type in Experiment 2 (approximately 1,800 ms), the N1000 component likely reflected the point at which participants were relating the difference comparison compound, analogically, to the difference sample compound. Insofar as this interpretation is accurate, it took about 1 s for participants to (covertly) relate difference relations.

Previous research in the neurocognitive literature has shown *negative* ERP components (e.g., N400) to be modulated by the ''cloze probability'' (i.e., degree of expectedness) of sentence-final words. For example, the sentence, ''it is hard to admit when one is *asleep*'' elicits a more negative waveform than the sentence ''it is hard to admit when one is *wrong*'' (Kutas, 1993; Kutas & Hillyard, 1984). Perhaps the negative waveforms elicited by different–different relational responding in Experiment 2 indicate that this analogical network (e.g., ''he is to his brother as chalk is to cheese'') is of low-probability relative to a similar–similar network (e.g., ''apple is to orange as dog is to cat''). In other words, low-probability/unusual relational responses may overlap functionally, to some extent, with responses to sentences that end with a low-probability word.

In Experiment 2, the bilateral prefrontal brain regions, especially the dorsolateral areas, were activated when participants engaged in the relating of derived similar–similar and different–different relations. Studies examining the neural substrates of natural-language analogy have found activity in the same regions (e.g., Luo et al., 2003). Interestingly, Experiment 2 showed ERP effects in the right hemisphere. Previous evidence suggests that right-hemispheric activation is, in some ways, special to analogical reasoning, especially analogies of the verbal kind (Luo et al., 2003). This evidence would seem to conflict with the well-established finding that linguistic processing, which clearly includes analogical reasoning, occurs predominantly in the left hemisphere. Certain verbal analogies may, however, have inductive, metaphorical, or ''open-ended'' qualities (Luo et al., 2003) leading to right-hemispheric prefrontal activa-

tion (Bottini et al., 1994; Goel & Dolan, 2000). Furthermore, deductive reasoning, in some ways the opposite of analogical reasoning, has yet to be associated with right-hemispheric frontal activation (Goel, Gold, Kapur, & Houle, 1998). Given the right-hemispheric activation found in Experiment 2, the current RFT model might have successfully captured this inductive or open-ended quality of verbal analogy. One issue that remains unresolved at the present time, however, concerns the fact that the difference between the similar–similar and different–different waveforms was reversed across the two hemispheres, although the right hemisphere difference only approached significance. Whether this lateral asymmetry also would be observed if a nonarbitrary relating-relations task (i.e., a figural analogy) were used remains to be seen (see Luo et al., 2003). In any case, until additional data are gathered with a larger sample, the current lateral asymmetry effect should be interpreted with caution.

The current research aimed to test the RFT model of analogical reasoning based on the relating of derived relations. The behavioral and neurophysiological data from Experiments 1 and 2 support the RFT prediction that similar–similar and different–different relational responses involve functionally distinct classes of behavior. The current research also shows that a behavior-analytic model of analogical reasoning can be generated in the laboratory, and that the model can be tested using chronometric and neurophysiological measures.

REFERENCES

Andreassi, J. L. (2000). *Psychophysiology: Human behavior & physiological response* (4th ed.). Hillsdale, NJ: Erlbaum.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44*, 91–124.

Barnes, D., Hegarty, N., & Smeets, P. M. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14*, 57–83.

Barnes-Holmes, D., Staunton, C., Barnes-Holmes, Y., Whelan, R., Stewart, I., Commins, S., et al. (2004). Interfacing relational frame theory with cognitive neuroscience: Semantic priming, the implicit association test, and event related potentials. *International Journal of Psychology and Psychological Therapy, 4*, 215–240.

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., et al. (2005). Derived stimulus relations, semantic priming, and event-related potentials: Testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior, 84*, 417–433.

Boroojerdi, B., Phipps, M., Kopylev, L., Wharton, C. M., Cohen, L. G., & Grafman, J. (2001). Enhancing analogic reasoning with rTMS over the left prefrontal cortex. *Neurology, 56*, 526–528.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., et al. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain, 117*, 1241–1253.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2002). Matching functionally same relations: Implications for equivalence-equivalence as a model for analogical reasoning. *The Psychological Record, 52*, 351–370.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record, 53*, 349–371.

Carpentier, F., Smeets, P. M., Barnes-Holmes, D., & Stewart, I. (2004). Matching derived functionally-same relations: Equivalence-equivalence and classical analogies. *The Psychological Record, 54*, 255–273.

Dymond, S., & Barnes, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior, 62*, 251–267.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64*, 163–184.

Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record, 46*, 271–300.

Goel, V., & Dolan, R. J. (2000). Anatomical segregation of component processes in an inductive inference task. *Journal of Cognitive Neuroscience, 12*, 110–121.

Goel, V., Gold, B., Kapur, S., & Houle, S. (1998). Neuroanatomical correlates of human reasoning. *Journal of Cognitive Neuroscience, 10*, 293–302.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language, 35*, 66–85.

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50*, 305–335.

Imam, A. A. (2001). Speed contingencies, number of stimulus presentations, and the nodality effect in equivalence class formation. *Journal of the Experimental Analysis of Behavior, 76*, 265–288.

Kastak, D., & Schusterman, R. J. (1994). Transfer of visual identity matching-to-sample in two California sea lions (Zalophus californianus). *Animal Learning & Behavior, 22*, 427–435.

Klein, G. A. (1987). Applications of analogical reasoning. *Metaphor and symbolic activity, 2*, 201–218.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8*, 533–572.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 1161–1163.

Lipkens, G. (1992). *Analogical reasoning as arbitrarily applicable relational responding.* Unpublished doctoral dissertation, University of Nevada, Reno.

Luo, Q., Perry, C., Peng, D., Jin, Z., Xu, D., Ding, G., & Xu, S. (2003). The neural substrate of analogical reasoning: An fMRI study. *Cognitive Brain Research, 17*, 527–534.

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56*, 519–555.

Stewart, I., & Barnes-Holmes, D. (2001). Understanding metaphor: A relational frame perspective. *The Behavior Analyst, 24*, 191–199.

Stewart, I., & Barnes-Holmes, D. (2004). A modern behavior-analytic approach to analogical reasoning. *The Irish Psychologist, 30*, 117–124.

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors and stories. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 73–86). New York: Plenum.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record, 51*, 381–408.

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2002). A functional-analytic model of analogy: A relational frame analysis. *Journal of the Experimental Analysis of Behavior, 78*, 275–396.

Weisbrod, M., Kefer, M., Winkler, S., Maier, S., Hill, H., Roesch-Ely, D., & Spitzer, M. (1999). Electrophysiological correlates of direct versus indirect semantic priming in normal volunteers. *Cognitive Brain Research, 8*, 289–298.

Wharton, C. M., Grafman, J., Flitman, S. S., Hansen, E. K., Brauner, J., Marks, A., & Honda, M. (2000). Towards neuroanatomical models of analogy: A positron emission tomography study of analogical mapping. *Cognitive Psychology, 40*, 173–197.

Whelan, R., & Barnes-Holmes, D. (2004). A derived transformation of consequential functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 82*, 177–195.

Received August 18, 2004
Final acceptance May 13, 2005